Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

Mukul Kumar Vishwas

Novosibirsk State University, Russia

Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros

27 April 2021

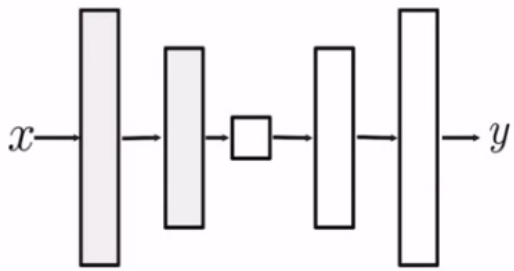
Problem statement

Image-to-Image translation is a class of CV and image processing where goal is to learn mapping between input and output image using a training set of align pairs. But most of the time pairs are not available or cost to make these pairs are too high cost and computation wise.

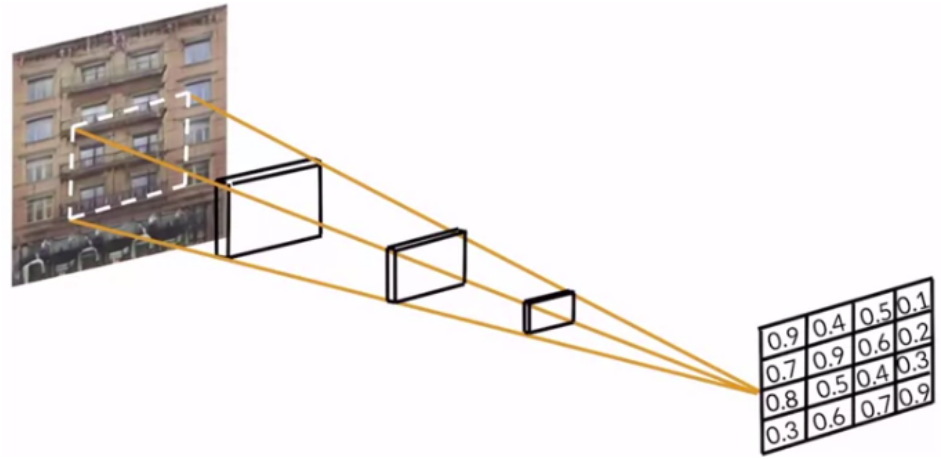
Solution: Learn a mapping G:X$\rightarrow$Y. such that distribution of X is indistinguishable from the Y using adversarial loss.

Prerequisite

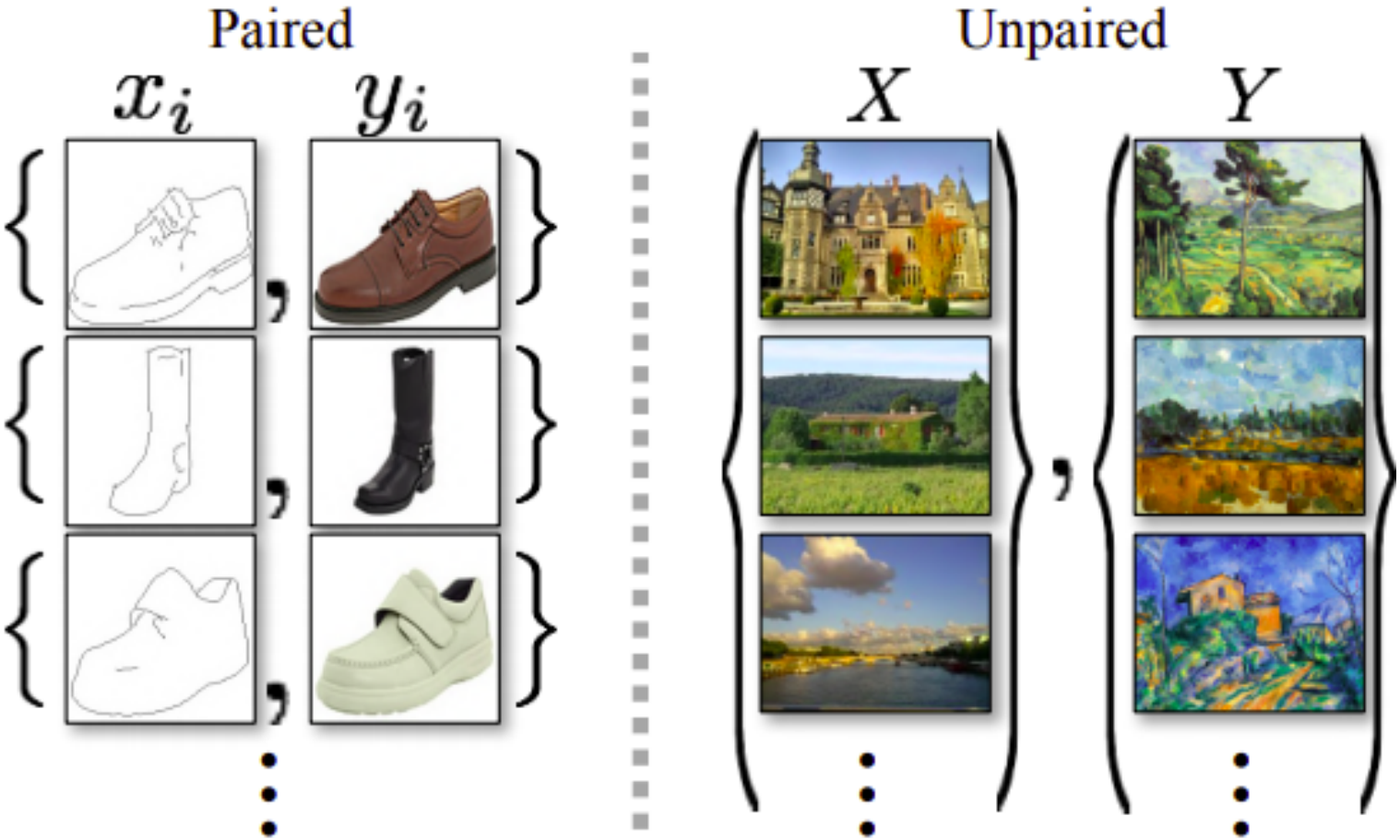
- What is GAN's and how they work.
- U-net Architecture (8 blocks).
  input size 256*256*3 encoded size 1*1*512



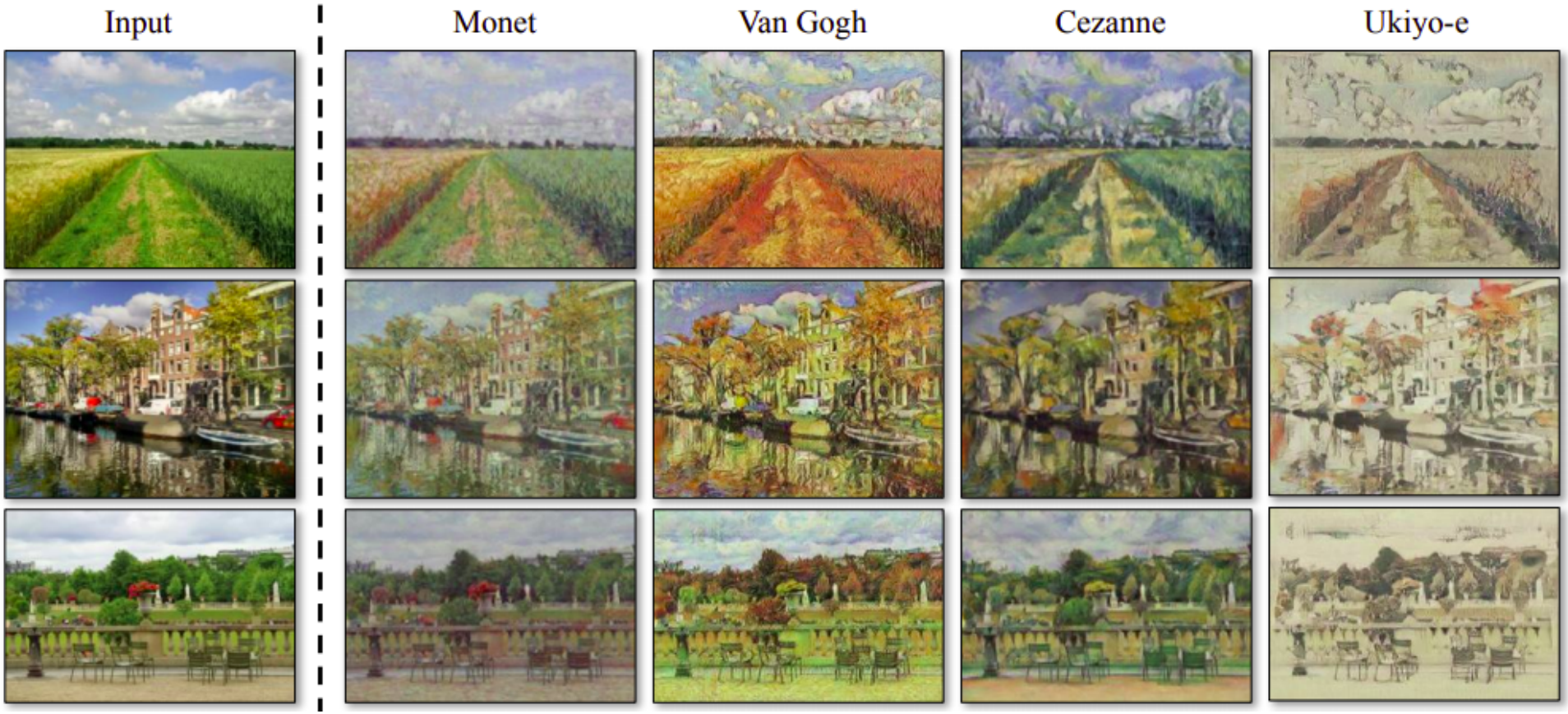
- Patch GAN.



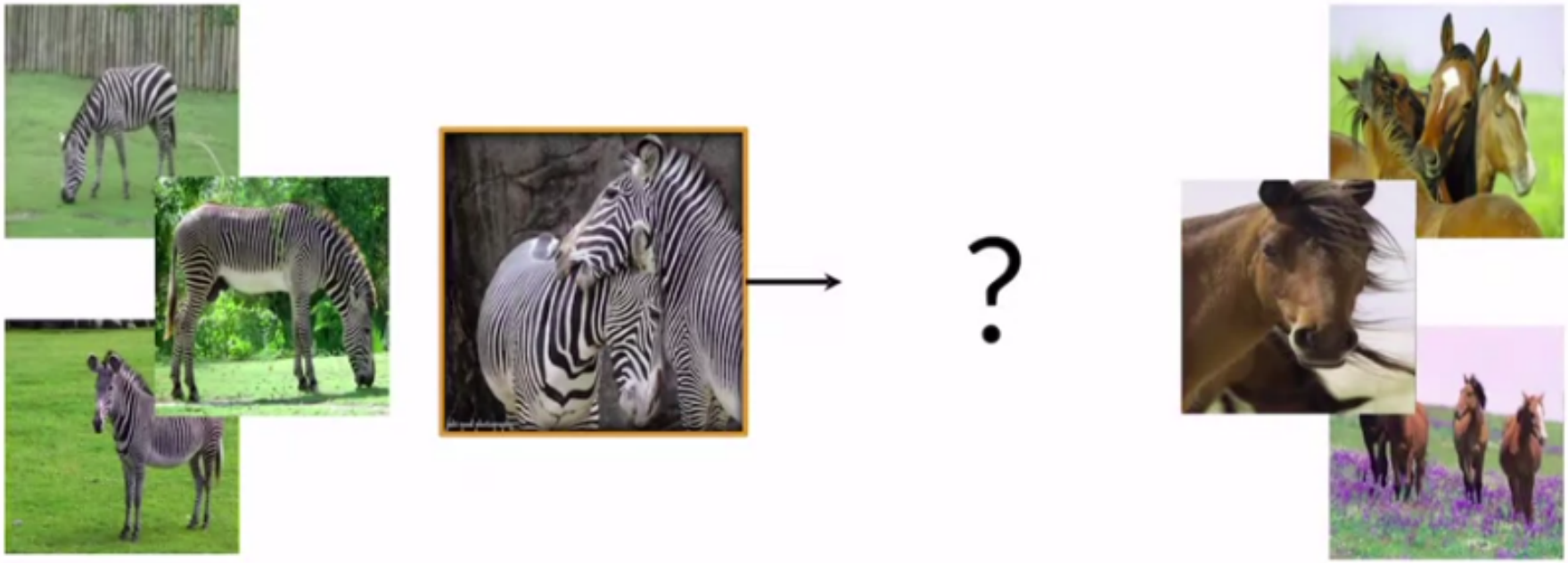
What is Paired and Unpaired data



What exactly is Image-to-Image translation

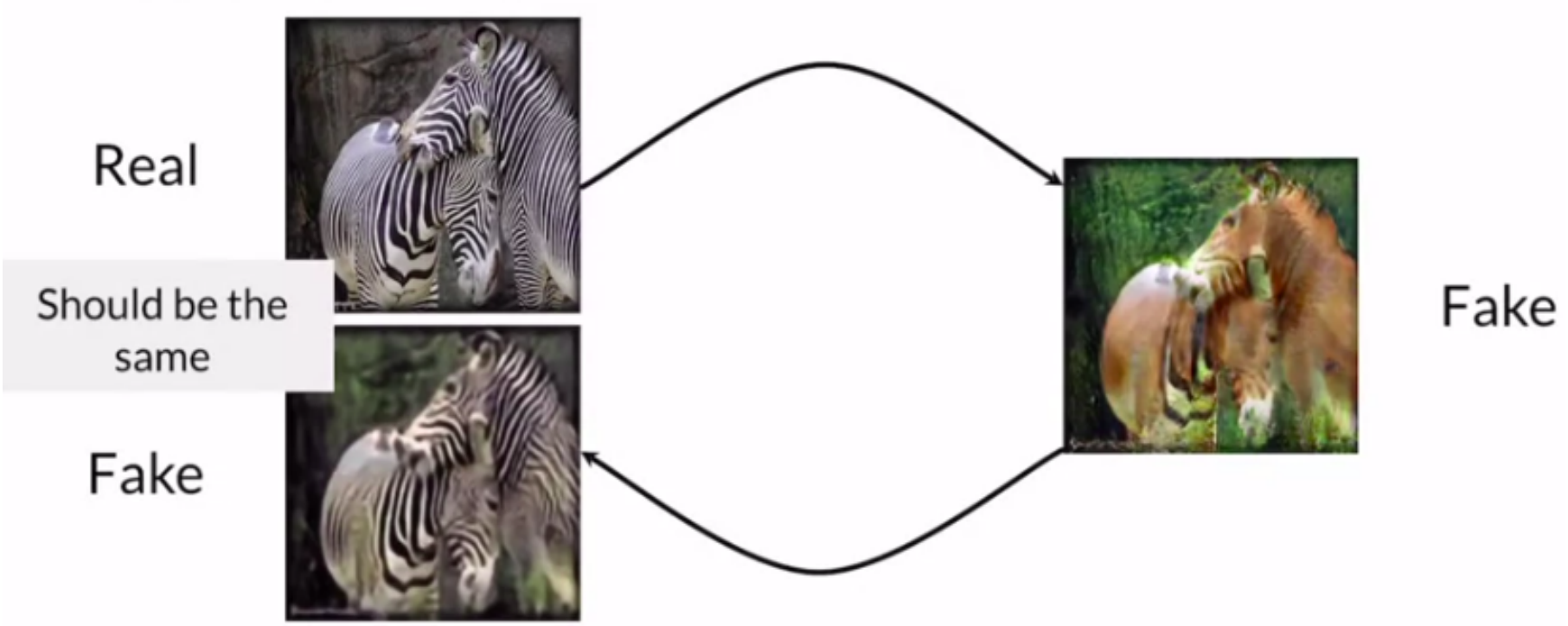


How it's possible



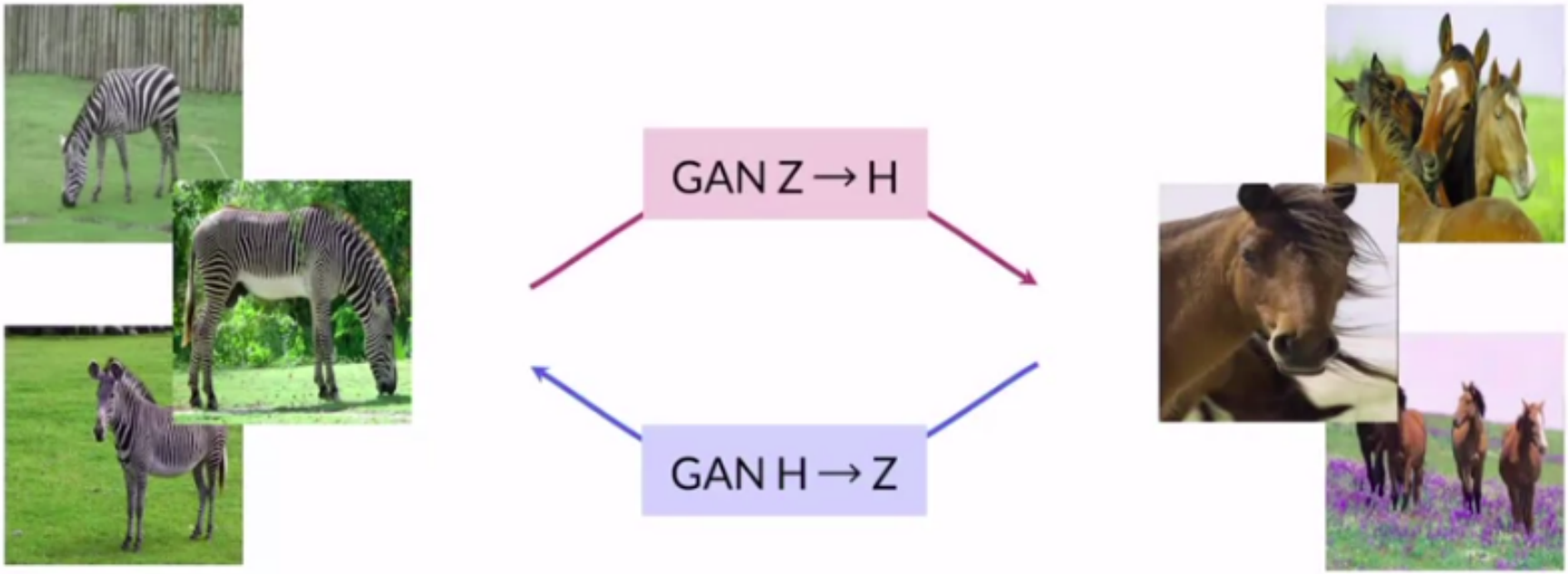
- Cycle GAN.
- What is cycle in Cycle GAN.
- How two GAN's working simultaneously.

Cycle consistency



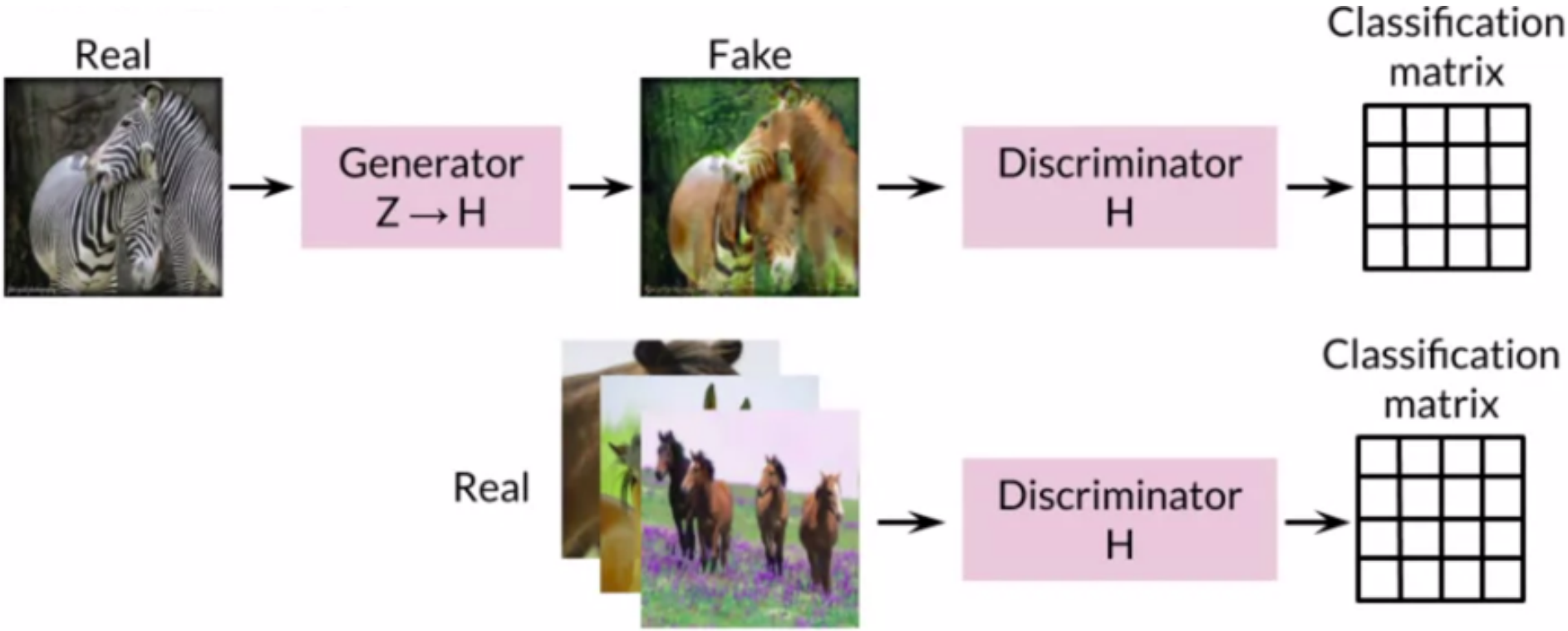
- Main idea is cycle consistency.
- Image flow from one style to another and back to original style.
- discriminator compare the real and final fake image.

Why 2 GAN's



- Adversarial and discriminator part responsible for realistic image generator.
- Cycle consistency responsible for content preservation.

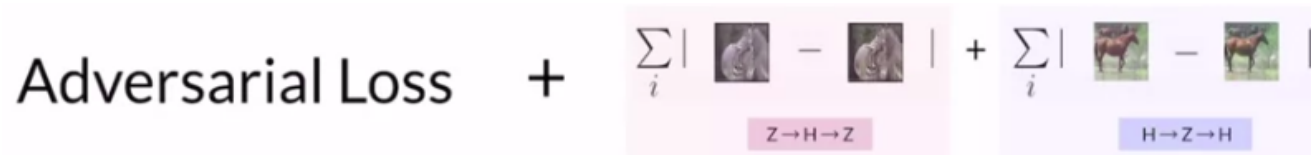
Architecture



- Zebra to horse generator.
- Discriminator distinguish between real or fake image.
- It's a patch GAN so output will be a classification matrix of patch in image.
- Same goes for backward direction.

Loss

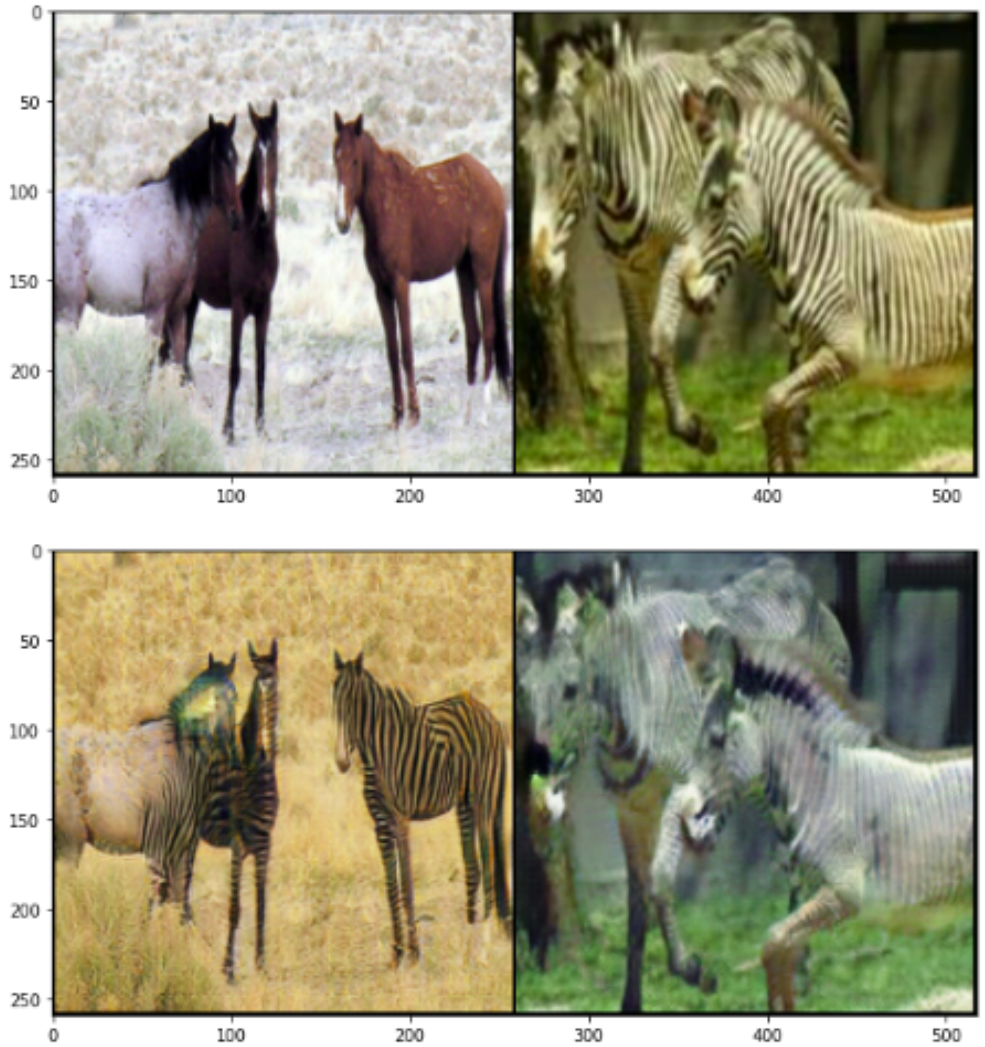
Total Loss = Adversarial + Cycle consistency



- Adversarial is MSE loss.
- Single optimizer for both generator.
- Adversarial loss is also divided into 2 parts A1(H2Z) and A2(Z2H).

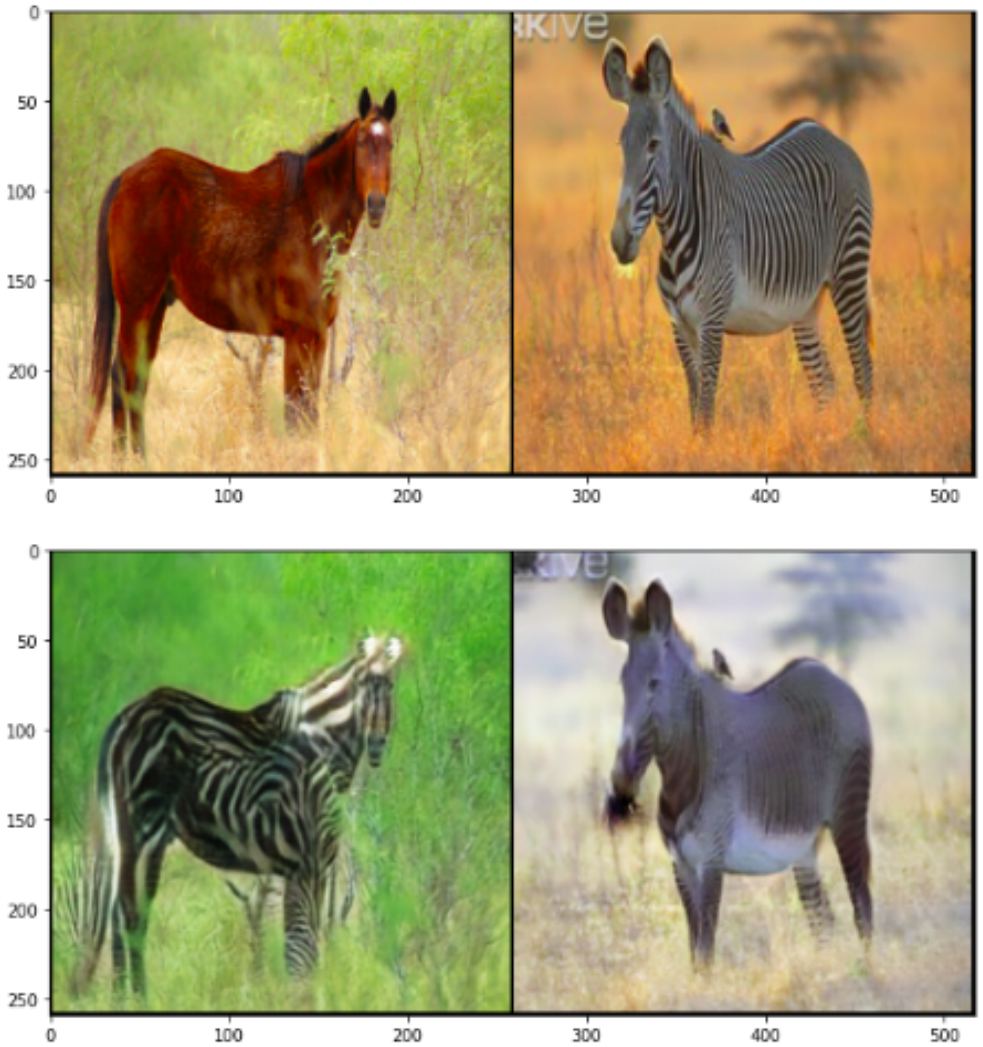
Experiment and Result

Epoch 120: Step 128800: Generator (U-Net) loss: 2.8694186446070695, Disc



Experiment and Result

Epoch 116: Step 124800: Generator (U-Net) loss: 2.9132890838384613, Dis



Thank you for your time.